

# An Extreme form of Superactivation for Quantum Zero-Error Capacities

Toby S. Cubitt and Graeme Smith

**Abstract**—The zero-error capacity of a channel is the rate at which it can send information perfectly, with zero probability of error, and has long been studied in classical information theory. We show that the zero-error capacity of quantum channels exhibits an extreme form of non-additivity, one which is not possible for classical channels, or even for the usual capacities of quantum channels. By combining probabilistic arguments with algebraic geometry, we prove that there exist channels  $\mathcal{E}_1$  and  $\mathcal{E}_2$  with no zero-error classical capacity whatsoever,  $C_0(\mathcal{E}_1) = C_0(\mathcal{E}_2) = 0$ , but whose joint zero-error quantum capacity is positive,  $Q_0(\mathcal{E}_1 \otimes \mathcal{E}_2) \geq 1$ . This striking effect is an extreme form of the *superactivation* phenomenon, as it implies that both the classical and quantum zero-error capacities of these channels can be superactivated simultaneously, whilst being a strictly stronger property of capacities. Superactivation of the quantum zero-error capacity was not previously known.

**Index Terms**—Additivity violation, channel coding, classical capacity, communication channels, information rates, quantum capacity, quantum theory, superactivation, zero-error capacity.

## I. INTRODUCTION

The zero-error capacity, introduced by Shannon in 1956, characterises the optimal achievable communication rate of a noisy channel when information must be transmitted with zero probability of error [1]. This is in contrast with the more traditional capacity, which only demands error probabilities vanishing in the limit of many channel uses. The question of zero-error capacity (and more generally zero-error information theory [2]) has a much more combinatorial flavor than the usual case, and has played an important role in the development of graph theory. Combinatorial optimisation problems are often intractable so, perhaps unsurprisingly, the zero-error capacity is unknown even for many very simple channels.

Quantum information theory seeks to extend information theory to include information sources and communication systems where quantum effects are important. Because all physical systems are fundamentally quantum, this can be seen as an attempt to more accurately model physical information

processing systems. Furthermore, expanding our notion of information to include quantum messages leads to new insights and applications, such as quantum cryptography and quantum computing. Because quantum systems are notoriously delicate, error correction is extremely important, and the capacities of a noisy quantum channel for transmitting various types of information noiselessly play a central role in the theory. In the context of zero-error quantum information theory, first studied in Ref. [3], the central capacities are the zero-error classical and zero-error quantum capacities.

A rather surprising effect has recently been discovered in the theory of quantum communication. Classically, there is a simple criterion for deciding whether a channel has non-zero capacity—any channel with some correlation between input and output has some positive capacity—and this criterion carries over to the usual classical capacity of quantum channels. However, when sending quantum information, the situation is very different. There are some quantum channels that are sufficiently noisy to have zero capacity for quantum communications, yet can still create correlations. In Ref. [4] it was shown that there are pairs of channels with very different noise characteristics, but both with zero quantum capacity, that, when used together, have a large joint quantum capacity. This *superactivation* is completely different from what happens in the classical case, and depends crucially on choosing entangled signal states for the joint channel.

Superactivation of classical channel capacities is easily seen to be impossible, both for the usual capacity and the zero-error capacity. If two classical channels have no correlation between input and output, so that their usual classical capacity vanishes, this will also hold for the joint channel. Similarly, if two classical channels each have the property that all pairs of inputs can lead to ambiguous outputs, so that the zero-error capacity vanishes, then the joint channel necessarily has this property too. The argument for the usual classical capacity carries over directly to the case of quantum channels; superactivation of the classical capacity of a quantum channel remains impossible.

However, in Ref. [5] it was shown that the *zero-error* classical capacity of a quantum channel actually *can* be superactivated (see also Ref. [6], which found superactivation of the non-asymptotic one-shot zero-error classical capacity, and a weaker form of activation in the asymptotic setting). In this paper, we significantly strengthen the results and techniques of Ref. [5]. There, techniques from algebraic geometry were combined with probabilistic arguments to show that there are pairs of channels, each with vanishing zero-error classical capacity, that have positive joint zero-error classical capacity

T. S. Cubitt is with the Departamento de Análisis Matemático, Universidad Complutense de Madrid, Plaza de Ciencias 3, Ciudad Universitaria, 28040 Madrid, Spain (email: tcubitt@mat.ucm.es).

G. Smith is with IBM T.J. Watson Research Center Yorktown Heights, NY 10598, USA (email: gbsmith@gmail.com)

T. S. Cubitt was supported by a Leverhulme early-career fellowship, and through the integrated EC project “QAP” (contract no. IST-2005-15848). G. Smith was supported by DARPA QUEST contract HR0011-09-C-0047. This work was carried out when T. S. Cubitt was with the Department of Mathematics, University of Bristol, United Kingdom

TSC would like to thank Nilanjana Datta and Francesco Buscemi for stimulating discussions about these ideas, and the physics of information group at IBM T. J. Watson Research Center for their hospitality during the visit in which this work was carried out.

when used together. Here, we find that there exist pairs of channels which each have vanishing zero-error classical capacity, as before, but when the two channels are used together they can even transmit more delicate *quantum* information with zero-error (indeed, only a single use of the joint channel is required). This is a particularly extreme form of superactivation, indeed it is the strongest possible form, and has not been seen previously for other capacities. It implies simultaneous superactivation of both the classical (already known from Ref. [5]) and quantum (previously unknown) zero-error capacities of quantum channels, whilst being strictly stronger than either of these.

The rest of the paper is organised as follows. In the next section we review some basic facts about quantum mechanics and algebraic geometry. Section III establishes sufficient conditions for this extreme form of superactivation, whilst Section IV shows that there exist channels which satisfy these conditions. Finally, Section V discusses the implications of our findings.

## II. PRELIMINARIES

### A. Quantum Mechanics

A minimum uncertainty state of a  $d$ -level quantum system is a pure state, represented by a  $d$ -dimensional complex unit vector  $|\psi\rangle \in \mathbb{C}^d$ . More generally, the state of a  $d$ -level system is given by a density matrix,  $\rho \in \mathcal{B}(\mathbb{C}^d)$ , where,  $\mathcal{B}(\mathbb{C}^d)$  denotes the set of bounded linear operators on  $\mathbb{C}^d$ . Such a density matrix is Hermitian ( $\rho = \rho^\dagger$ ) and has unit trace,  $\text{Tr } \rho = 1$ . As a result, any such  $\rho$  admits a spectral decomposition  $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$  with orthogonal  $|\psi_i\rangle$ , which can be interpreted as describing a system that is in state  $|\psi_i\rangle$  with probability  $p_i$ . Whilst we will not need to consider measurement processes below, we will need to know when there is some measurement to perfectly distinguish two states. This is possible exactly when the states are orthogonal, i.e. for pure states when  $\langle\psi|\varphi\rangle = 0$ , or for mixed states when  $\text{Tr } \rho\sigma = 0$ .

It is sometimes useful to consider (unnormalised) pure states  $|\psi\rangle_{AB}$  in a bipartite space  $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$  as matrices  $M = \mathbb{M}(|\psi\rangle_{AB})$  in the isomorphic space of  $d_A \times d_B$  matrices  $\mathcal{M}_{d_A, d_B}$ . The isomorphism arises from fixing some product basis  $|i\rangle_A |j\rangle_B$  for  $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ , and expanding  $|\psi\rangle_{AB} = \sum_{i,j} M_{ij} |i\rangle_A |j\rangle_B$  in this basis. A bipartite subspace  $S \subseteq \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$  is isomorphic in this way to a matrix subspace which we denote  $\mathbb{M}(S)$ .

We define the “flip” operation on a bipartite state as the operation that swaps the two systems and takes the complex conjugate:

$$\mathbb{F}(|\psi\rangle_{AB}) = \text{SWAP}(|\bar{\psi}\rangle_{AB}). \quad (1)$$

In terms of the matrix representation  $M = \mathbb{M}(|\psi\rangle_{AB})$ , the flip operation is just Hermitian conjugation:  $\mathbb{M}(\mathbb{F}|\psi\rangle_{AB}) = M^\dagger$ . The flip operation can be extended to operators and subspaces in the obvious way.

The most general physical operation in quantum mechanics is a completely-positive trace preserving (CPT) map from  $\mathcal{B}(\mathbb{C}^{d_{\text{in}}})$  to  $\mathcal{B}(\mathbb{C}^{d_{\text{out}}})$ , where  $d_{\text{in}}$  and  $d_{\text{out}}$  are the input and output dimensions of the map. We will refer to such

operations as *quantum channels* throughout, as they are directly analogous to channels in classical information theory. A quantum channel that maps a space  $\mathcal{H}_A$  to  $\mathcal{H}_B$  can always be thought of as an isometry followed by a partial trace. In other words, for any channel  $\mathcal{E}$  we have  $\mathcal{E}(\rho) = \text{Tr}_E U \rho U^\dagger$ , where  $U : \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_E$  is an isometry satisfying  $U^\dagger U = I_A$ . Equivalently, the action of a channel can be expressed in terms of Kraus operators:  $\mathcal{N}(\rho) = \sum_k A_k \rho A_k^\dagger$ , where  $\sum_k A_k^\dagger A_k = I_A$ . A third representation of quantum channels (indeed, it extends to any linear map), which plays an important role in Ref. [5], is the *Choi-Jamiołkowski matrix*, defined to be the result of applying the channel to one half of an unnormalised maximally entangled state. In other words, the Choi-Jamiołkowski matrix of a channel  $\mathcal{E}$  is given by  $\sigma = (\mathcal{I} \otimes \mathcal{E})(\omega)$  where  $|\omega\rangle = \sum_{i=1}^{d_A} |i\rangle |i\rangle$ , and  $\omega = |\omega\rangle\langle\omega|$ . The action of the channel can be recovered from the Choi-Jamiołkowski matrix via  $\mathcal{E}(\rho) = \text{Tr}_A [\sigma_{AB} \cdot \rho_A^T \otimes \mathbb{1}_B]$  (where  $\rho_A^T$  denotes the transpose of the density matrix  $\rho_A$ ).

We will also need the adjoint  $\mathcal{E}^*$  of a channel  $\mathcal{E}$ , which is simply the dual with respect to the Hilbert-Schmidt inner product, i.e. the unique map defined by:

$$\text{Tr}[X^\dagger \mathcal{E}(Y)] = \text{Tr}[\mathcal{E}^*(X)^\dagger Y]. \quad (2)$$

In terms of Kraus operators  $A_k$ , the adjoint  $\mathcal{E}^*$  of  $\mathcal{E}$  is the map whose Kraus operators are the Hermitian conjugates  $A_k^\dagger$ . (Note that  $\mathcal{E}^*$  is CP, but not necessarily trace-preserving.)

### B. Algebraic Geometry

In order to prove our results, we need some basic notions from algebraic geometry (see e.g. Ref. [7]). A key concept is that of a *Zariski-closed* set, and the resulting *Zariski topology*. We will only ever work over base fields  $\mathbb{C}$  or  $\mathbb{R}$ , so for our purposes Zariski-closed sets are sets defined by a collection of polynomials, i.e. they are the solution sets of simultaneous polynomial equations. We will use the terms *Zariski-closed set* and *algebraic set* interchangeably.

The Zariski topology is the topology whose *closed* sets are the Zariski-closed sets. It is the standard topology in algebraic geometry, but it serves more as a convenient terminology than providing any useful geometric information. The main use we will make of it is the fact that intersections of Zariski-closed sets are themselves Zariski-closed. Indeed, the only Zariski-closed set that has non-zero measure (in the usual sense on  $\mathbb{C}^d$  or  $\mathbb{R}^d$ ) is the entire space. This “Zariski dichotomy”—that a Zariski-closed set is either zero-measure or the entire space—lies at the heart of our proofs.

We will also frequently refer to the *Grassmannian*  $\text{Gr}_d(V)$  of a vector space  $V$ , the set of all  $d$ -dimensional subspaces of  $V$ . There is a standard way of embedding the Grassmannian in projective space, called the *Plücker embedding* and conventionally denoted  $\iota$ . If a  $d$ -dimensional subspace in the Grassmannian is spanned by some basis  $\{|\psi_i\rangle\}$ , then  $\iota(S)$  is defined to be  $\wedge_{i=1}^d |\psi_i\rangle$ , with  $\wedge$  denoting the anti-symmetric product. This is uniquely defined, since picking some other basis replaces  $|\psi_i\rangle$  by  $\sum_{j=1}^d A_{i,j} |\psi_j\rangle$  for some invertible matrix  $A$ , which in turn replaces  $\iota(S)$  by  $\det(A)\iota(S)$ . In

projective space, rescaling by the scalar  $\det(A)$  makes no difference.

Via the Plücker embedding, points in the Grassmannian are naturally parametrised by the coordinates of points in projective space, called the *Plücker coordinates*. (Note that not all points in the ambient projective space correspond to points in the Grassmannian; the Plücker coordinates of points within the Grassmannian must satisfy quadratic constraints called the *Plücker relations*.) Thus the Plücker coordinates  $P_{\alpha_d}$  of  $S$  are defined by  $\sum_{\alpha_d} P_{\alpha_d} (\wedge_{j \in \alpha_d} |j\rangle) = \wedge_{i=1}^d |\psi_i\rangle$ , where  $\alpha_d$  are size  $d$  subsets of  $\{1 \dots n\}$ , with  $|1\rangle, \dots, |n\rangle$  a basis of  $V$ .

### III. SUFFICIENT CONDITIONS FOR SUPERACTIVATION

We start by reducing the problem of proving existence of our extreme form of superactivation to a question about the existence of subspaces satisfying certain conditions. The arguments are very similar to those leading to Theorem 13 of Ref. [5], but the stronger requirement that the joint channel have positive *quantum* zero-error capacity adds an additional constrain on the subspaces. To derive this new constraint, we need the following lemma, which gives us a sufficient condition for a channel to have positive zero-error quantum capacity.

**Lemma 1** *Let  $\mathcal{E} : \mathcal{H}_A \rightarrow \mathcal{H}_B$  be a channel,  $|0\rangle$  and  $|1\rangle$  be states on  $\mathcal{H}_A$ , and  $|\pm\rangle = 1/\sqrt{2}(|0\rangle \pm |1\rangle)$ . Then, if*

$$\text{Tr}[\mathcal{E}(|0\rangle\langle 0|) \mathcal{E}(|1\rangle\langle 1|)] = 0 \quad (3)$$

and

$$\text{Tr}[\mathcal{E}(|+\rangle\langle +|) \mathcal{E}(|-\rangle\langle -|)] = 0, \quad (4)$$

we have  $Q_0(\mathcal{E}) \geq 1$ .

*Proof:* To see this, suppose  $A_k$  are the Kraus operators of  $\mathcal{E}$  and  $\varphi = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|)$ , and let

$$\mathcal{R}_\varphi(\rho) = \sum_k \sqrt{\varphi} A_k^\dagger \mathcal{E}(\varphi)^{-1/2} \rho \mathcal{E}(\varphi)^{-1/2} A_k \sqrt{\varphi} + \Pi \rho \Pi, \quad (5)$$

where  $\mathcal{E}(\varphi)^{-1/2}$  is the square-root of the Moore-Penrose pseudo-inverse of  $\mathcal{E}(\varphi)$  (i.e. its inverse when restricted to its support), and  $\Pi$  is the projector onto the kernel of  $\mathcal{E}(\varphi)$  (which vanishes if  $\mathcal{E}(\varphi)$  is invertible). This corresponds to the reversal operation of Ref. [8] when  $\mathcal{E}(\varphi)$  is full rank. It is completely-positive and trace preserving by design, and  $\mathcal{M} := \mathcal{R}_\varphi \circ \mathcal{E}$  is the identity on  $\text{span}(|0\rangle, |1\rangle)$ .

To see this, first note that, by assumption,

$$0 = \text{Tr}[\mathcal{E}(|0\rangle\langle 0|) \mathcal{E}(|1\rangle\langle 1|)] \quad (6a)$$

$$= \sum_{j,k} \text{Tr} \left[ A_j |0\rangle\langle 0| A_j^\dagger A_k |1\rangle\langle 1| A_k^\dagger \right] \quad (6b)$$

$$= \sum_{j,k} \left| \langle 0| A_j^\dagger A_k |1\rangle \right|^2, \quad (6c)$$

so that  $\langle 0| A_j^\dagger A_k |1\rangle = 0$  for all  $j, k$ , and similarly for  $\langle +| A_j^\dagger A_k |-\rangle$ . Now consider

$$\text{Tr} \left[ \sqrt{\varphi} A_k^\dagger \mathcal{E}(\varphi)^{-1/2} \mathcal{E}(|0\rangle\langle 0|) \mathcal{E}(\varphi)^{-1/2} A_k \sqrt{\varphi} |1\rangle\langle 1| \right] \quad (7a)$$

$$= \frac{1}{2} \text{Tr} \left[ A_k^\dagger \mathcal{E}(\varphi)^{-1/2} \mathcal{E}(|0\rangle\langle 0|) \mathcal{E}(\varphi)^{-1/2} A_k |1\rangle\langle 1| \right] \quad (7b)$$

$$= \frac{1}{2} \text{Tr} \left[ \mathcal{E}(\varphi)^{-1/2} \mathcal{E}(|0\rangle\langle 0|) \mathcal{E}(\varphi)^{-1/2} \mathcal{E}(|1\rangle\langle 1|) \right]. \quad (7c)$$

Since  $\mathcal{E}(|0\rangle\langle 0|)$  and  $\mathcal{E}(|1\rangle\langle 1|)$  are orthogonal, we have

$$\mathcal{E}(\varphi)^{-1/2} = \sqrt{2} \mathcal{E}(|0\rangle\langle 0|)^{-1/2} + \sqrt{2} \mathcal{E}(|1\rangle\langle 1|)^{-1/2}, \quad (8)$$

which immediately implies with Eq. (7) and the fact that  $\mathcal{E}(|0\rangle\langle 0|)$  and  $\mathcal{E}(|1\rangle\langle 1|)$  are in the support of  $\mathcal{E}(\varphi)$  that

$$\mathcal{M}(|0\rangle\langle 0|) = |0\rangle\langle 0| \quad (9a)$$

$$\mathcal{M}(|1\rangle\langle 1|) = |1\rangle\langle 1|. \quad (9b)$$

Similarly, we also have

$$\mathcal{M}(|+\rangle\langle +|) = |+\rangle\langle +| \quad (10a)$$

$$\mathcal{M}(|-\rangle\langle -|) = |-\rangle\langle -|. \quad (10b)$$

Now all we have to do is show that any CPT map  $\mathcal{M}$  satisfying the above four equations must be the identity. We can easily use these four equations to show that

$$\mathcal{M}(\mathbb{I}) = \mathbb{I}, \quad (11a)$$

$$\mathcal{M}(X) = X, \quad (11b)$$

$$\mathcal{M}(Z) = Z, \quad (11c)$$

where  $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$  and  $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ .

Since  $\mathcal{M}$  is a unital qubit channel [9], it is a mixture of conjugations of Pauli matrices of the form

$$\mathcal{M}(\rho) = (1 - p_X - p_Y - p_Z) \rho + p_X X \rho X + p_Y Y \rho Y + p_Z Z \rho Z, \quad (12)$$

where  $Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ . This form, together with Eq. (11a), implies that  $p_X = p_Y = p_Z = 0$ , so that  $\mathcal{M}(\rho) = \rho$ . ■

We are now in a position to reduce our superactivation problem to a question about subspaces. The approach is the closely related to that in Ref. [5], which in turn builds on the techniques of Ref. [10]. We start by recapping the conditions required for superactivation of the *classical* zero-error capacity from Ref. [5], which is necessary (but not sufficient) for our result. We then show how to strengthen this to achieve the extreme form of superactivation claimed here.

Recall that two quantum states  $\rho, \sigma$  are perfectly distinguishable if and only if they are orthogonal ( $\text{Tr}[\rho\sigma] = 0$ ). Thus, the *classical* zero-error capacity of a channel  $\mathcal{E}$  is 0 iff no pair of inputs gives orthogonal outputs:

$$\forall \psi, \varphi : \quad 0 \neq \text{Tr}[\mathcal{E}(\varphi) \mathcal{E}(\psi)] = \text{Tr}[\psi \cdot \mathcal{E}^* \circ \mathcal{E}(\varphi)], \quad (13)$$

where we have simply pulled the channel across the inner product in the final equality, giving the composition of the adjoint  $\mathcal{E}^*$  and the channel. Rewriting these expressions by expressing the action of the composite map  $\mathcal{N} = \mathcal{E}^* \circ \mathcal{E}$  in terms of its Choi-Jamiołkowski state  $\sigma$ , this is equivalent to:

$$\forall \psi, \varphi : \quad \text{Tr}[\sigma \cdot \varphi_A^T \otimes \psi_{A'}] \neq 0. \quad (14)$$



But this simply expresses the condition that  $\sigma$  should not be orthogonal to any product state. Therefore, for the channel  $\mathcal{E}$  to have no classical zero-error capacity, the support  $S$  of  $\sigma$  must contain no product states in its orthogonal complement  $S^\perp$ .

Thus in order to superactivate the *one-shot, classical* zero-error capacity, we need two subspaces  $S_1, S_2$  (corresponding to two channels  $\mathcal{E}_1, \mathcal{E}_2$  as described above), each of which has no product states in its orthogonal complement, such that the joint channel  $\mathcal{E}_1 \otimes \mathcal{E}_2$  *does* have positive classical zero-error capacity. To achieve superactivation even in the asymptotic setting, we must strengthen the condition on the individual subspaces to ensure that even arbitrarily many copies of the individual channels have no capacity. Since the Choi-Jamiołkowski matrix of  $k$  copies  $\mathcal{N}^{\otimes k}$  of a map is the tensor power  $\sigma^{\otimes k}$  of the single-copy Choi-Jamiołkowski matrix, this is equivalent to requiring that no *tensor power*  $S_{1,2}^{\otimes k}$  of either subspace has a product state in its orthogonal complement.

As well as the individual channels having no capacity, we also want the joint channel  $\mathcal{E}_1 \otimes \mathcal{E}_2$  to have positive zero-error capacity, i.e. we require the converse of Eq. (13) to hold for the joint channel:

$$\exists \psi, \varphi : \text{Tr}[\psi \cdot (\mathcal{E}_1^* \circ \mathcal{E}_1) \otimes (\mathcal{E}_2^* \circ \mathcal{E}_2)(\varphi)] = 0. \quad (15)$$

Let us choose the (unnormalised) inputs  $|\psi\rangle, |\varphi\rangle$  to the joint channel to be the maximally entangled states  $|\omega\rangle$  and  $\mathbb{1} \otimes X |\omega\rangle$ , where  $X$  is now the generalisation of the Pauli  $X$  matrix to arbitrary dimension, i.e. the matrix with ones down its anti-diagonal. Expressing Eq. (15) in terms of the Choi-Jamiołkowski matrix  $\sigma_1 \otimes \sigma_2$  of the joint channel, the condition of Eq. (15) simplifies to:

$$\text{Tr}[\sigma_1^T \cdot (\mathbb{1} \otimes X) \sigma_2 (\mathbb{1} \otimes X^\dagger)] = 0. \quad (16)$$

This simply expresses the condition that  $(X \otimes \mathbb{1}) \sigma_2 (X^\dagger \otimes \mathbb{1})$  and  $\sigma_1^T$  should have orthogonal supports, i.e.  $(X \otimes \mathbb{1}) S_2 \perp S_1^T$ . Since we also want the individual subspaces to have no product states in their orthogonal complements, it makes sense to choose the two subspaces to be as big as possible (so that their orthogonal complements are as small as possible), subject to this condition. We therefore choose  $S_2$  to be the orthogonal complement (up to the local unitary rotation and transposition) of  $S_1$ :

$$S_2^T = \mathbb{1} \otimes X \cdot S_1^\perp. \quad (17)$$

This allows us to express all the requirements for classical zero-error superactivation in terms of conditions on a *single* subspace  $S := S_1$ . These conditions are summarised in the following theorem (which is Theorem 13 from Ref. [5]):

**Theorem 2** *If there exists a subspace  $S$  and unitaries  $U, V$  satisfying*

$$\forall k, \nexists |\psi\rangle, |\varphi\rangle \in \mathcal{H}_A^{\otimes k} : |\psi\rangle \otimes |\varphi\rangle \in (S^{\otimes k})^\perp, \quad (18a)$$

$$\forall k, \nexists |\psi\rangle, |\varphi\rangle \in \mathcal{H}_A^{\otimes k} : |\psi\rangle \otimes |\varphi\rangle \in ((S^\perp)^{\otimes k})^\perp, \quad (18b)$$

$$\mathbb{F}(S) = S, \quad (18c)$$

$$\mathbb{F}(\mathbb{1} \otimes X \cdot S) = \mathbb{1} \otimes X \cdot S, \quad (18d)$$

$$\exists \{M_i \geq 0\} : \mathbb{M}(S) = \text{span}\{M_i\}, \quad (18e)$$

$$\exists \{M_j \geq 0\} : \mathbb{M}(\mathbb{1} \otimes X \cdot S^\perp) = \text{span}\{M_j\}, \quad (18f)$$

*then there exist channels  $\mathcal{E}_{1,2}$  with  $C_0(\mathcal{E}_1) = C_0(\mathcal{E}_2) = 0$  but  $C_0(\mathcal{E}_1 \otimes \mathcal{E}_2) \geq 1$ .*

The final four conditions in Eqs. (18c) to (18f) express the requirement that the subspace must come from the support of a Choi-Jamiołkowski matrix of a composite map with the very particular form  $\mathcal{E}^* \circ \mathcal{E}$ , which imposes additional symmetries on the subspace. (Fuller details of the proof can be found in Ref. [5, Theorem 13].)

Theorem 2 gives sufficient conditions for superactivation of the *classical* zero-error capacity. But we want something significantly stronger; we not only want the joint channel to have positive classical zero-error capacity, we want it even to have positive *quantum* zero-error capacity. For this, we must strengthen Theorem 2 using Lemma 1:

**Theorem 3** *Suppose there is a subspace  $S$  of a bipartite Hilbert space  $\mathcal{H}_A \otimes \mathcal{H}_A$  such that*

$$\forall k, \nexists |\psi\rangle, |\varphi\rangle \in \mathcal{H}_A^{\otimes k} : |\psi\rangle \otimes |\varphi\rangle \in (S^{\otimes k})^\perp, \quad (19a)$$

$$\forall k, \nexists |\psi\rangle, |\varphi\rangle \in \mathcal{H}_A^{\otimes k} : |\psi\rangle \otimes |\varphi\rangle \in ((S^\perp)^{\otimes k})^\perp, \quad (19b)$$

$$\mathbb{F}(S) = S, \quad (19c)$$

$$\mathbb{F}(\mathbb{1} \otimes X \cdot S) = \mathbb{1} \otimes X \cdot S, \quad (19d)$$

$$\exists \{M_i \geq 0\} : \mathbb{M}(S) = \text{span}\{M_i\}, \quad (19e)$$

$$\exists \{M_j \geq 0\} : \mathbb{M}(\mathbb{1} \otimes X \cdot S^\perp) = \text{span}\{M_j\}, \quad (19f)$$

$$S \perp (\mathbb{1} + X) \otimes (\mathbb{1} - X) S^\perp. \quad (19g)$$

*Then there exist channels  $\mathcal{E}_{1,2}$  with  $C_0(\mathcal{E}_1) = C_0(\mathcal{E}_2) = 0$  but  $Q_0(\mathcal{E}_1 \otimes \mathcal{E}_2) \geq 1$ .*

*Proof:* Equations (19a) to (19f) are identical to the conditions in Theorem 2, and already give sufficient conditions for the individual channels to have no zero-error capacity,  $C_0(\mathcal{E}_1) = C_0(\mathcal{E}_2) = 0$ , and the joint channel to have positive *classical* zero-error capacity,  $C_0(\mathcal{E}_1 \otimes \mathcal{E}_2) \geq 1$ . Only Eq. (19g) is new. We must show that this additional condition is sufficient to ensure the joint channel has positive *quantum* zero-error capacity,  $Q_0(\mathcal{E}_1 \otimes \mathcal{E}_2) \geq 1$ .

Recall from Theorem 2 and Ref. [5] that  $S$  will be the support of  $\sigma_1 = (\mathcal{I} \otimes \mathcal{E}_1^* \circ \mathcal{E}_1)(\omega)$  with  $|\omega\rangle = \sum_i |i\rangle |i\rangle$  and  $S_2 = (\mathbb{1} \otimes X) S^\perp$  the support of  $\sigma_2^T$  defined similarly. The two signal states for  $\mathcal{E}_1 \otimes \mathcal{E}_2$  in Theorem 2 are  $|\varphi_0\rangle = |\omega\rangle$  and  $|\varphi_1\rangle = (\mathbb{1} \otimes X) |\omega\rangle$ . From Lemma 1, what we have to do now is show that, letting  $|\varphi_\pm\rangle = (|\varphi_0\rangle \pm |\varphi_1\rangle)/\sqrt{2}$ , we have

$$\text{Tr}[(\mathcal{E}_1 \otimes \mathcal{E}_2)(|\varphi_+\rangle\langle\varphi_+|)(\mathcal{E}_1 \otimes \mathcal{E}_2)(|\varphi_-\rangle\langle\varphi_-|)] = 0. \quad (20)$$

Now,

$$\text{Tr}[(\mathcal{E}_1 \otimes \mathcal{E}_2)(|\varphi_+\rangle\langle\varphi_+|)(\mathcal{E}_1 \otimes \mathcal{E}_2)(|\varphi_-\rangle\langle\varphi_-|)] \quad (21a)$$

$$= \text{Tr}[(\mathcal{E}_1^* \circ \mathcal{E}_1 \otimes \mathcal{E}_2^* \circ \mathcal{E}_2)(|\varphi_+\rangle\langle\varphi_+|)|\varphi_-\rangle\langle\varphi_-|] \quad (21b)$$

$$= \text{Tr} \left[ \sigma_1^{A'_1 A_1} \otimes \sigma_2^{A'_2 A_2} |\varphi_+\rangle\langle\varphi_+|_{A'_1 A'_2}^T \otimes |\varphi_-\rangle\langle\varphi_-|_{A_1 A_2} \right] \quad (21c)$$

$$= \text{Tr} \left[ \sigma_1^{A'_1 A_1} \otimes \sigma_2^{A'_2 A_2} \left( P_+ \otimes \mathbb{1} |\omega\rangle\langle\omega|_{A'_1 A'_2} P_+ \otimes \mathbb{1} \right) \otimes \left( P_- \otimes \mathbb{1} |\omega\rangle\langle\omega|_{A_1 A_2} P_- \otimes \mathbb{1} \right) \right] \quad (21d)$$

$$= \text{Tr} \left[ (P_+ \otimes P_-) \sigma_1^{A'_1 A_1} (P_+ \otimes P_-) \otimes \sigma_2^{A'_2 A_2} \cdot \right. \\ \left. |\omega\rangle\langle\omega|_{A'_1 A'_2} \otimes |\omega\rangle\langle\omega|_{A_1 A_2} \right] \quad (21e)$$

$$= \text{Tr} \left[ [(P_+ \otimes P_-) \sigma_2 (P_+ \otimes P_-)]^T \sigma_1 \right] \quad (21f)$$

where  $P_\pm = (\mathbb{1} \pm X)/2$  are projectors and we have used the fact that  $\text{Tr}[|\omega\rangle\langle\omega| M \otimes N] = \text{Tr}[N^T M]$ . As a result, the requirement Eq. (20) is met by choosing  $S \perp (P_+ \otimes P_-)S_2$ . This is equivalent to  $S \perp (\mathbb{1} + X) \otimes (\mathbb{1} - X)S^\perp$ , since  $P_- X = -P_-$  and we chose  $S_2 = (\mathbb{1} \otimes X)S^\perp$ . ■

#### IV. EXISTENCE OF SUPERACTIVATION

Given Theorem 3, all we need to do in order to show the extreme superactivation phenomenon is to prove that there do exist subspaces satisfying the conditions of the theorem. We use a combination of algebraic-geometry and probabilistic arguments to establish this result.

In what follows, we will need to consider a number of sets of subspaces. Recall the definition of extendibility from Refs. [5], [11]:

**Definition 4** A subspace  $S \subseteq \mathcal{H}_A \otimes \mathcal{H}_B$  is  $k$ -unextendible if  $(S^{\otimes k})^\perp$  contains no product state in  $\mathcal{H}_{A^{\otimes k}} \otimes \mathcal{H}_{B^{\otimes k}}$ . A subspace is strongly unextendible if it is  $k$ -unextendible for all  $k \geq 1$ . Conversely, a subspace is  $k$ -extendible if it is not  $k$ -unextendible, and extendible if it is not strongly unextendible.

Following Ref. [5], we denote the sets of  $d$ -dimensional  $k$ -extendible, extendible, and unextendible subspaces, respectively, by

$$E_d^k(\mathcal{H}_A, \mathcal{H}_B) = \{S \in \text{Gr}_d(\mathcal{H}_A \otimes \mathcal{H}_B) | S \text{ is } k\text{-extendible}\}, \quad (22)$$

$$E_d(\mathcal{H}_A, \mathcal{H}_B) = \{S \in \text{Gr}_d(\mathcal{H}_A \otimes \mathcal{H}_B) | S \text{ is extendible}\}, \quad (23)$$

$$U_d(\mathcal{H}_A, \mathcal{H}_B) = \{S \in \text{Gr}_d(\mathcal{H}_A \otimes \mathcal{H}_B) | S \text{ is unextendible}\}, \quad (24)$$

Note that the set  $\bigcup_k E_d^k$  is the set of subspaces that do contain product states in their orthogonal complements, so it is precisely the set of subspaces that we want to avoid in order to satisfy the condition in Eq. (19a). At the heart of our proof is the following Lemma, which shows that the set  $E_d^k$  algebraic:

**Lemma 5**  $E_d^k(\mathcal{H}_A, \mathcal{H}_B)$  is Zariski-closed in  $\text{Gr}_d(\mathcal{H}_A \otimes \mathcal{H}_B) = \text{Gr}_d(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ .

This is proven in Lemma 15 of Ref. [5] using standard algebraic geometry arguments, based on the fact that there is a simple algebraic characterisation of product states  $|\psi\rangle_{AB}$  as those states for which  $\mathbb{M}(|\psi\rangle_{AB})$  is rank 1.

We will also refer to the set

$$F_d(\mathbb{R}, d_A) = \{S \in \text{Gr}_{2d}(\mathbb{R}^2 \otimes \mathbb{R}^{d_A} \otimes \mathbb{R}^{d_A}) | \\ S = iS, \mathbb{F}(S) = S, \mathbb{F}(\mathbb{1} \otimes X \cdot S) = \mathbb{1} \otimes X \cdot S\} \quad (25)$$

of subspaces satisfying the symmetry constraints of Eqs. (19c) and (19d). Note that we are considering  $F_d$  as a subset of the real Grassmannian, in which context  $i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ .

**Lemma 6**  $F_d(\mathbb{R}, d_A)$  is Zariski-closed in  $\text{Gr}_{2d}(\mathbb{R}^2 \otimes \mathbb{R}^{d_A} \otimes \mathbb{R}^{d_A})$ .

This is proven in Lemma 17 of Ref. [5], writing out the constraints on  $S$  from Eq. (25) explicitly in terms of the Plücker coordinates, and verifying that the constraints are polynomials.

In order to extend the arguments of Ref. [5] to our case, we will need to consider an additional set: the set of subspaces satisfying the orthogonality constraint of Eq. (19g):

$$C_d(\mathbb{C}, d_A) = \{S \in \text{Gr}_d(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_A}) \\ | S \perp (\mathbb{1} + X) \otimes (\mathbb{1} - X)S^\perp\}, \quad (26)$$

and also the isomorphic set of real vector spaces:

$$C_d(\mathbb{R}, d_A) = \{S \in \text{Gr}_{2d}(\mathbb{R}^2 \otimes \mathbb{R}^{d_A} \otimes \mathbb{R}^{d_A}) \\ | S = iS, S \perp (\mathbb{1} + X) \otimes (\mathbb{1} - X)S^\perp\}. \quad (27)$$

The first step is to show that this set is algebraic (cf. Lemma 17 of Ref. [5]).

**Lemma 7**  $C_d(\mathbb{C}, d_A)$  is Zariski-closed in  $\text{Gr}_d(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_A})$ .

*Proof:* First, we let  $W = \bigwedge_{i=1}^d |\psi_i\rangle$  for some basis  $\{|\psi_i\rangle\}$  of  $S$ . We have  $|\psi\rangle \in S$  exactly when  $|\psi\rangle \wedge W = 0$  and we want to use this to construct a basis for  $S^\perp$ . If  $P_{\alpha_d}$  are the Plücker coordinates of  $S$ , and supposing  $|\psi\rangle = \sum_{i=1}^n v_i |i\rangle$ , then  $|\psi\rangle$  is in  $S$  exactly when

$$|\psi\rangle \wedge W = \sum_{\alpha_d} \sum_i v_i P_{\alpha_d} |i\rangle \wedge (\bigwedge_{j \in \alpha_d} |j\rangle) \quad (28a)$$

$$= \sum_{i, \beta_{d+1}} v_i N_{i, \beta_{d+1}} \bigwedge_{k \in \beta_{d+1}} |k\rangle = 0, \quad (28b)$$

so that we have an  $N$  such that  $|\psi\rangle \in S$  iff  $\langle\psi| N = 0$ . Now, the support of  $NN^\dagger$  is  $S^\perp$  and its eigenvalues are positive. Most importantly, we can think of  $NN^\dagger$  as a matrix with entries that are quadratic polynomials in  $P_{\alpha_d}$ . Thus, we are interested in ensuring that

$$N \cdot P_+ \otimes P_- |\psi\rangle = 0 \quad (29)$$

for all  $|\psi\rangle \in S$ , which is equivalent to showing that

$$N^{\otimes n} (P_+ \otimes P_-)^{\otimes n} \iota(S) = 0. \quad (30)$$

This is a linear constraint on  $\iota(S)$ , so  $\{\iota(S) : N^{\otimes n}(P_+ \otimes P_-)^{\otimes n} \iota(S) = 0\}$  is Zariski-closed. Since  $\iota$  is a proper morphism (cf. Lemma 17 of Ref. [5]),  $C_d = \{S : \forall |\psi\rangle \in S, N \cdot P_+ \otimes P_- |\psi\rangle = 0\}$  must also be Zariski-closed. ■

Any Zariski-closed set in a complex vector space is also Zariski-closed in the isomorphic real vector space. Furthermore, the intersection of two Zariski-closed sets is again Zariski-closed, since they form a topology. This immediately gives:

**Corollary 8**  $E_d^k(\mathcal{H}_A, \mathcal{H}_{A'}) \cap F_d(\mathbb{R}, d_A) \cap C_d(\mathbb{R}, d_A)$  is Zariski-closed in  $F_d(\mathbb{R}, d_A) \cap C_d(\mathbb{R}, d_A)$ .

We can now use the ‘‘Zariski dichotomy’’ to prove that the set of strongly unextendible subspaces is full measure in  $F_d \cap C_d$ . Note that our results are not particularly sensitive to the choice of measure, but for definiteness, when we refer to a measure or to a probability distribution on the Grassmannian, this can always be taken to be the one induced by the Haar measure over the unitary group. More explicitly, the action of the unitary group on a Hilbert space induces a natural measure on quantum states — the standard choice in quantum information theory. This in turn induces a measure on subspaces of a given dimension, i.e. on the Grassmannian. When refer to a measure on a subset of the Grassmannian, we mean the restriction of the measure over the whole Grassmannian to that subset.

We will make use of *unextendible product bases* in the proofs, which are defined as follows:

**Definition 9** An unextendible product basis (UPB) is a set of product states  $\{|\psi_i\rangle_{AB}\}$  (not necessarily orthogonal) in a bipartite space  $\mathcal{H}_A \otimes \mathcal{H}_B$  such that  $(\text{span}\{|\psi_i\rangle\})^\perp$  contains no product states. The dimension of a UPB is the number of product states in the set.

Clearly, a UPB spans a 1-unextendible subspace. In fact, Lemma 22 of Ref. [5], which we restate here, shows that the span is even strongly unextendible:

**Lemma 10** If  $\{|\psi_i^1\rangle_{A_1B_1}\}$  and  $\{|\psi_i^2\rangle_{A_2B_2}\}$  are unextendible product bases in  $\mathcal{H}_{A_1} \otimes \mathcal{H}_{B_1}$  and  $\mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2}$  respectively, then  $\{|\psi_i^1\rangle|\psi_j^2\rangle\}_{i,j}$  is an unextendible product basis in  $\mathcal{H}_{A_1A_2} \otimes \mathcal{H}_{B_1B_2}$ .

We are now in a position to prove the following key lemma.

**Lemma 11** For  $d \geq 12(d_A + d_B - 1)$ , the set of strongly unextendible subspaces  $U_d(\mathcal{H}_A, \mathcal{H}_{A'}) \cap F_d(\mathbb{C}, d_A) \cap C_d(\mathbb{C}, d_A)$  is full measure in  $F_d(\mathbb{C}, d_A) \cap C_d(\mathbb{C}, d_A)$ .

*Proof:* Since  $E_d^k(\mathcal{H}_A, \mathcal{H}_{A'}) \cap F_d(\mathbb{R}, d_A) \cap C_d(\mathbb{R}, d_A)$  is Zariski-closed by Corollary 8,  $\bigcup_k E_d^k(\mathcal{H}_A, \mathcal{H}_{A'}) \cap F_d(\mathbb{R}, d_A) \cap C_d(\mathbb{R}, d_A)$  is a countable union of Zariski-closed sets, so it is either zero measure in  $F_d(\mathbb{R}, d_A) \cap C_d(\mathbb{R}, d_A)$ , or it is the full space. Conversely, its complement  $U_d(\mathcal{H}_A, \mathcal{H}_{A'}) \cap F_d(\mathbb{R}, d_A) \cap C_d(\mathbb{R}, d_A)$  is either full measure or empty.

To rule out the possibility that it is empty, we prove that there exists a subspace in  $U_d \cap F_d \cap C_d$  by constructing one using unextendible product bases (UPBs). Lemma 10 shows that the span of a UPB is a strongly unextendible subspace,

and it is known from Ref. [12] that UPBs of dimension  $m$  exist in  $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$  for any  $m \geq d_A + d_B - 1$ . Let  $S$  be a subspace spanned by such a minimal UPB, and let the set of matrices  $\{M_i\}$  be a basis for  $\mathbb{M}(S)$ . Consider the symmetrised subspace  $\mathbb{M}(S')$  spanned by

$$\begin{aligned} &\{M, XMX, M^\dagger, XM^\dagger X, \\ &P_+MP_-, P_+XMXP_-, P_+M^\dagger P_-, P_+XM^\dagger XP_-, \\ &P_-MP_+, P_-XMXP_+, P_-M^\dagger P_+, P_-XM^\dagger XP_+\}. \end{aligned} \quad (31)$$

The resulting subspace  $S'$  has dimension at most  $12(d_A + d_B - 1)$ , and satisfies both the symmetry and orthogonality constraints of Eqs. (19c), (19d) and (19g) from Theorem 3. Thus  $S' \in F_d \cap C_d$ . Since  $S$  is strongly-unextendible, and  $S \subseteq S'$ ,  $S'$  is clearly strongly unextendible, which completes the proof. ■

**Corollary 12** For any  $d_A \geq 48$ , and for a subspace  $S \in \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_A}$  of dimension  $12(2d_A - 1) \leq d \leq d_A^2 - 12(2d_A - 1)$  chosen at random<sup>1</sup> subject to the constraints  $\mathbb{F}(S) = S$ ,  $\mathbb{F}(\mathbb{1} \otimes X \cdot S) = \mathbb{1} \otimes X \cdot S$  and  $S \perp (\mathbb{1} + X) \otimes (\mathbb{1} - X) S^\perp$ , both  $S$  and  $S^\perp$  will almost-surely be strongly unextendible.

*Proof:* Lemma 11 implies that  $S$  chosen in this way will almost-surely be strongly unextendible. But  $S^\perp$  is then a random subspace subject to the same constraints, with dimension  $12(2d_A - 1) \leq d^\perp = d_A^2 - d \leq d_A^2 - 12(2d_A - 1)$ . Thus Lemma 11 implies that  $S^\perp$  will also be almost-surely strongly unextendible. For there to exist a suitable  $d$ , we require  $12(2d_A - 1) \leq d_A^2 - 12(2d_A - 1)$ , or  $d_A \geq 48$ . ■

Corollary 8 tells us that, although Eqs. (19a) and (19b) of Theorem 3 would appear to impose severe constraints on the subspace  $S$ , they are in fact benign. Even if we restrict to subspaces satisfying Eqs. (19c), (19d) and (19g), a randomly chosen subspace will satisfy Eqs. (19a) and (19b) with probability 1.

It remains to show that such a subspace can also satisfy Eqs. (19e) and (19f). For this, we require more information about the structure of the set  $F_d \cap C_d$  of subspaces that simultaneously satisfy Eqs. (19c), (19d) and (19g).

**Lemma 13** If  $d_A$  is even, then

$$\begin{aligned} &F_d(\mathbb{R}, d_A) \cap C_d(\mathbb{R}, d_A) \\ &\cong \bigsqcup_{r=\max[0, d-\frac{d_A^2}{2}]}^{\min[d, \frac{d_A^2}{2}]} \bigsqcup_{k_1=0}^r \bigsqcup_{k_2=0}^{d-r} \left( \text{Gr}_{k_1}(\mathbb{R}^{d_A^2/2}) \times \right. \\ &\quad \left. \text{Gr}_{r-k_1}(\mathbb{R}^{d_A^2/2}) \times \text{Gr}_{k_2}(\mathbb{R}^{d_A^2/2}) \times \text{Gr}_{d-r-k_2}(\mathbb{R}^{d_A^2/2}) \right). \end{aligned} \quad (32)$$

The  $\sqcup$  denotes disjoint union, meaning an element of  $F_d(\mathbb{R}, d_A) \cap C_d(\mathbb{R}, d_A)$  can be uniquely identified by specifying non-negative integers  $r$ ,  $k_1$  and  $k_2$  satisfying  $d - d_A^2/2 \leq r \leq d$ ,  $k_1 \leq r$  and  $k_2 \leq d - r$ , along with elements of  $\text{Gr}_{k_1}(\mathbb{R}^{d_A^2/2})$ ,  $\text{Gr}_{r-k_1}(\mathbb{R}^{d_A^2/2})$ ,  $\text{Gr}_{k_2}(\mathbb{R}^{d_A^2/2})$  and  $\text{Gr}_{d-r-k_2}(\mathbb{R}^{d_A^2/2})$ .

<sup>1</sup>E.g. according to the distribution induced by the Haar measure; see discussion preceding Lemma 11.

*Proof:* Elements of  $F_d(\mathbb{R}, d_A) \cap C_d(\mathbb{R}, d_A)$  are  $2d$ -dimensional real subspaces of  $\mathbb{R}^2 \otimes \mathbb{R}^{d_A} \otimes \mathbb{R}^{d_A}$ . As such, they can be expressed as rank- $2d$  projectors. In terms of these projectors  $\Pi$ , the constraints in Eq. (25) defining  $F_d(\mathbb{R}, d_A)$  become  $i \Pi i^T = \Pi$ ,  $\mathbb{F} \Pi \mathbb{F}^T = \Pi$  and  $(X \otimes X) \Pi (X \otimes X) = \Pi$  (cf. Lemma 28 of Ref. [5]).

The additional constraint  $S \perp (P_+ \otimes P_-) S^\perp$  in Eq. (27) defining  $C_d(\mathbb{R}, d_A)$  can also be expressed as a symmetry of  $\Pi$ . Note that this constraint is symmetric; if  $S$  satisfies it, then so does  $S^\perp$ . To see this, express the constraint as

$$\forall |\psi\rangle \in S^\perp, |\varphi\rangle \in S : \langle \varphi | P_+ \otimes P_- | \psi \rangle = 0 \quad (33)$$

and take the complex conjugate. If  $\Pi$  is the projector corresponding to a subspace  $S$ , the constraint is equivalent to

$$\Pi(P_+ \otimes P_-) \Pi = (P_+ \otimes P_-) \Pi, \quad (34)$$

and we know the same holds for  $S^\perp$ :

$$(\mathbb{1} - \Pi)P(\mathbb{1} - \Pi) = P(\mathbb{1} - \Pi), \quad (35)$$

or, equivalently,

$$\Pi(P_+ \otimes P_-) \Pi = \Pi(P_+ \otimes P_-). \quad (36)$$

Together, Eqs. (34) and (36) imply that if Eq. (34) is satisfied then  $\Pi$  and  $P_+ \otimes P_-$  commute. Conversely, it is easy to see that Eq. (34) is satisfied if  $\Pi$  commutes with  $P_+ \otimes P_-$ . Thus the subspace  $S$  is in  $C_d(\mathbb{R}, d_A)$  iff  $\Pi$  commutes with  $(P_+ \otimes P_-) = (\mathbb{1} + X) \otimes (\mathbb{1} - X)$  and  $i \Pi i^T = \Pi$ .

We will first consider the  $P_+ \otimes P_-$  and  $\mathbb{F}$  symmetries. Since  $\Pi$  commutes with  $P_+ \otimes P_-$ , it must be of the form  $\Pi = \Pi_{P_{+-}} + \Pi_{P_{+-}}^\perp$  where  $\Pi_{P_{+-}}$  is a projector onto a subspace in the support of  $P_+ \otimes P_-$ , and  $\Pi_{P_{+-}}^\perp$  is a projector onto a subspace in the orthogonal complement thereof. Note that, as we are working in the real vector space,  $P_+ \otimes P_-$  is rank  $d_A^2/2$ . Now,  $\mathbb{F}$  exchanges  $P_+ \otimes P_-$  with  $P_- \otimes P_+$ , so  $\mathbb{F} \Pi \mathbb{F}^T = \Pi_{P_{+-}} + \Pi_{P_{+-}}^\perp$ , where  $\Pi_{P_{+-}}$  ( $\Pi_{P_{+-}}^\perp$ ) is a projector onto a subspace in the (orthogonal complement of the) support of  $P_- \otimes P_+$ . But  $\mathbb{F} \Pi \mathbb{F}^T = \Pi$ , so  $\Pi_{P_{+-}}$  must commute with  $\Pi_{P_{+-}}^\perp$  and, furthermore,  $\Pi_{P_{+-}} = \mathbb{F} \Pi_{P_{+-}} \mathbb{F}^T$ . Thus

$$\Pi = (\Pi_{P_{+-}} + \mathbb{F} \Pi_{P_{+-}} \mathbb{F}^T) + \Pi^\perp, \quad (37)$$

where  $\Pi^\perp$  is a projector onto a subspace in the support of  $\mathbb{1} - (P_+ \otimes P_- + P_- \otimes P_+)$  that satisfies  $\mathbb{F} \Pi^\perp \mathbb{F}^T = \Pi^\perp$ . Let  $r \leq d$  denote the rank of  $\Pi_{P_{+-}}$ . Since  $P_+ \otimes P_-$  has  $d_A^2/2$  dimensional support,  $r$  cannot be larger than this. Also, as  $\Pi$  has rank  $2d$ ,  $\Pi^\perp$  has rank  $2(d-r)$ . But  $\Pi^\perp$  must live in the support of  $\mathbb{1} - (P_+ \otimes P_- + P_- \otimes P_+)$  which has dimension  $d_A^2$ , so we require  $2(d-r) \leq d_A^2$ . Thus  $r$  is constrained to take values in the range

$$\max[0, d - d_A^2/2] \leq r \leq \min[d, d_A^2/2]. \quad (38)$$

Now consider the  $i$  and  $\mathbb{F}$  symmetries. Since  $P_+ \otimes P_- + P_- \otimes P_+$  is invariant under both these operations,  $\Pi_{P_{+-}} + \mathbb{F} \Pi_{P_{+-}} \mathbb{F}^T$  and  $\Pi^\perp$  must satisfy these symmetries independently. We first focus on  $\Pi^\perp$ . Let  $\mathbb{F}_\pm$  denote the  $\pm 1$  eigenspaces of  $\mathbb{F}$ . Since  $\Pi^\perp$  commutes with  $\mathbb{F}$ , it must be the sum of a projector onto a subspace of  $\mathbb{F}_+$  and a projector onto a subspace of  $\mathbb{F}_-$ . In other words,  $\Pi^\perp = \Pi_+^\perp + \Pi_-^\perp$  where

$\Pi_\pm \mathbb{F} = \mathbb{F} \Pi_\pm = \pm \Pi_\pm$ . Since  $i$  and  $\mathbb{F}$  anti-commute,  $i$  must map  $\mathbb{F}_\pm$  to  $\mathbb{F}_\mp$ . Thus  $i \Pi_\pm^\perp i^T$  is a projector onto  $\mathbb{F}_\mp$ . Combined with the fact that  $i \Pi^\perp i^T = \Pi^\perp$  we obtain  $i \Pi_\pm^\perp i^T = \Pi_\mp^\perp$ . We can thus assume that

$$\Pi^\perp = \Pi_+^\perp + i \Pi_+^\perp i^T \quad (39)$$

where  $\Pi_+^\perp$  is a projector onto  $\mathbb{F}_+$  within the support of  $\mathbb{1} - (P_+ \otimes P_- + P_- \otimes P_+)$ . Since  $\Pi^\perp$  has rank  $2(d-r)$ ,  $\Pi_+^\perp$  must have rank  $d-r$ .

Turning now to  $\Pi_{P_{+-}} + \mathbb{F} \Pi_{P_{+-}} \mathbb{F}^T$ , this already commutes with  $\mathbb{F}$ , so we must be able to rewrite it as  $\Pi_{P_{+-}} + \mathbb{F} \Pi_{P_{+-}} \mathbb{F}^T = \Pi_+ + \Pi_-$  where  $\Pi_\pm$  are projectors onto  $\mathbb{F}_\pm$  within the support of  $P_+ \otimes P_- + P_- \otimes P_+$ . By the same argument as before, the  $i$  symmetry imposes  $\Pi_- = i \Pi_+ i^T$ , so

$$\Pi_{P_{+-}} + \mathbb{F} \Pi_{P_{+-}} \mathbb{F}^T = \Pi_+ + i \Pi_+ i^T. \quad (40)$$

Since  $\Pi_{P_{+-}} + \mathbb{F} \Pi_{P_{+-}} \mathbb{F}^T$  has rank  $2r$ ,  $\Pi_+$  must have rank  $r$ .

Finally, consider the  $X \otimes X$  symmetry. Since  $X \otimes X$  commutes with  $\mathbb{F}$  and  $P_{+-} \otimes P_{+-}$ , we have that  $\Pi_+$  and  $\Pi_+^\perp$  must also commute with  $X \otimes X$ . This means we can write  $\Pi_+$  as  $\Pi_{++} + \Pi_{+-}$  and  $\Pi_+^\perp$  as  $\Pi_{++}^\perp + \Pi_{+-}^\perp$ , where  $\Pi_{\pm\pm}$ ,  $\Pi_{\pm\pm}^\perp$  are projectors onto subspaces of the  $\pm 1$  eigenspace of  $X \otimes X$ . Since  $\Pi_+$  has rank  $r$ , the ranks of  $\Pi_{++}$  and  $\Pi_{+-}$  must sum to  $r$ . Similarly,  $\Pi_+^\perp$  has rank  $d-r$ , so the ranks of  $\Pi_{++}^\perp$  and  $\Pi_{+-}^\perp$  must sum to  $d-r$ . Thus we have finally that

$$\begin{aligned} \Pi &= \Pi_{++} + \Pi_{+-} + \Pi_{++}^\perp + \Pi_{+-}^\perp \\ &\quad + i(\Pi_{++} + \Pi_{+-} + \Pi_{++}^\perp + \Pi_{+-}^\perp)i^T. \end{aligned} \quad (41)$$

Conversely, if  $\Pi_{++}$ ,  $\Pi_{+-}$ ,  $\Pi_{++}^\perp$  and  $\Pi_{+-}^\perp$  are arbitrary projectors with the appropriate supports and with ranks summing to  $r$  and  $d-r$ , respectively, then a  $\Pi$  of the above form projects onto a subspace in  $F_d(\mathbb{R}, d_A) \cap C_d(\mathbb{R}, d_A)$ . For each value of  $r$  satisfying Eq. (38), if  $\Pi_{++}$  and  $\Pi_{++}^\perp$  have ranks  $k_1$  and  $k_2$ , then our choice of  $\Pi$  is equivalent to choosing an element of  $\text{Gr}_{k_1}(\mathbb{R}^{d_A^2/2}) \times \text{Gr}_{r-k_1}(\mathbb{R}^{d_A^2/2}) \times \text{Gr}_{k_2}(\mathbb{R}^{d_A^2/2}) \times \text{Gr}_{d-r-k_2}(\mathbb{R}^{d_A^2/2})$ . ■

This structure lemma allows us to deal with the remaining conditions of Theorem 3, namely Eqs. (19e) and (19f), using probabilistic arguments.

**Theorem 14** *If  $d_A$  is even, and  $\lfloor d/2 \rfloor \leq d_A^2/2 - 2$ , then the set*

$$\begin{aligned} P_d(d_A) &= \{S \in F_d(\mathbb{C}, d_A) \cap C_d(\mathbb{C}, d_A) \\ &\quad | \exists M \in \mathbb{M}(S), M' \in \mathbb{M}(\mathbb{1} \otimes X \cdot S^\perp) : M, M' \geq 0\} \end{aligned} \quad (42)$$

*has non-zero measure in  $F_d(\mathbb{C}, d_A) \cap C_d(\mathbb{C}, d_A)$ .*



*Proof:* Since  $\dim \text{Gr}_k(\mathbb{R}^{d_A^2/2}) = (d_A^2/2 - k)k$ , we have

$$\begin{aligned} & \dim \left( \text{Gr}_{k_1}(\mathbb{R}^{d_A^2/2}) \times \text{Gr}_{r-k_1}(\mathbb{R}^{d_A^2/2}) \right. \\ & \quad \left. \times \text{Gr}_{k_2}(\mathbb{R}^{d_A^2/2}) \times \text{Gr}_{d-r-k_2}(\mathbb{R}^{d_A^2/2}) \right) \\ &= \left( \frac{d_A^2}{2} - k_1 \right) k_1 \left( \frac{d_A^2}{2} - r + k_1 \right) (r - k_1) \\ & \quad \cdot \left( \frac{d_A^2}{2} - k_2 \right) k_2 \left( \frac{d_A^2}{2} - d + r + k_1 \right) (d - r - k_2), \end{aligned} \quad (43)$$

which takes its maximum value at  $r = d/2$ ,  $k_1 = k_2 = d/4$  for  $d$  a multiple of 4, or the closest integers to this otherwise. This means that all but a measure-zero subset of  $F_d(\mathbb{C}, d_A) \cap C_d(\mathbb{C}, d_A)$  is contained in the component associated with these values of  $r$ ,  $k_1$  and  $k_2$ . Indeed, if  $d$  is a multiple of 4 then the component of  $F_d(\mathbb{C}, d_A) \cap C_d(\mathbb{C}, d_A)$  corresponding to  $\text{Gr}_{d/4}(\mathbb{R}^{d_A^2/2}) \times \text{Gr}_{d/4}(\mathbb{R}^{d_A^2/2}) \times \text{Gr}_{d/4}(\mathbb{R}^{d_A^2/2}) \times \text{Gr}_{d/4}(\mathbb{R}^{d_A^2/2})$  has measure 1 in  $F_d(\mathbb{C}, d_A)$ . Otherwise, the components corresponding to the closest integers to  $r = d/2$ ,  $k_1 = k_2 = d/4$  together have total measure 1, with the measure split equally between them. For the remainder of the proof we will take  $r = d/2$ ,  $k_1 = k_2 = d/4$  ( $d$  divisible by 4) or any set of closest integers to these. Let  $K_d(\mathbb{C}, d_A)$  denote the corresponding part of  $F(\mathbb{C}, d_A) \cap C_d(\mathbb{C}, d_A)$ .

It suffices to show that  $P_d(d_A) \cap K_d(\mathbb{C}, d_A)$  has positive measure in  $K_d(\mathbb{C}, d_A)$ . To do so, we first construct a subspace  $S \in K_d(\mathbb{C}, d_A)$  that contains a positive-definite element (i.e.  $M > 0$  for some  $M \in \mathbb{M}(S)$ ), such that  $(\mathbb{1} \otimes X)S^\perp$  also contains a positive-definite element. This will guarantee that every  $S' \in K_d(\mathbb{C}, d_A)$  that is sufficiently close to  $S$  will contain a positive-semidefinite element, hence will belong to  $P_d(d_A) \cap K_d(\mathbb{C}, d_A)$ , implying that this set has non-zero measure and proving the theorem.

To construct the desired  $S$ , choose  $S$  to contain  $|\omega\rangle = \sum_{i=1}^{d_A} |i, i\rangle$ , which has  $M = \|(|\omega\rangle)\| = \mathbb{1} \geq 0$ . We will also require that  $S$  be orthogonal to  $(\mathbb{1} \otimes X)|\omega\rangle$  so that  $(\mathbb{1} \otimes X)S^\perp$  also contains  $|\omega\rangle$  and is positive definite. (Note that this only works if  $d_A$  is even, otherwise  $|\omega\rangle$  and  $(\mathbb{1} \otimes X)|\omega\rangle$  are not orthogonal.)

$P_\pm \otimes P_\mp |\omega\rangle = P_\pm \otimes P_\mp (\mathbb{1} \otimes X)|\omega\rangle = 0$ , so both  $|\omega\rangle$  and  $(\mathbb{1} \otimes X)|\omega\rangle$  are contained in the support of  $\mathbb{1} - (P_+ \otimes P_- + P_- \otimes P_+)$ . They also both belong to the  $+1$  eigenspace of  $X \otimes X$ . Thus to choose  $S$  we need only choose an additional  $k_2 - 1$  dimensions for  $\Pi_{++}^\perp$  (from a space of dimension  $d^2/4 - 1$ ) as well as an arbitrary rank- $(d - r - k_2)$  projector  $\Pi_{+-}^\perp$  whose support is contained within the portion of the  $-1$  eigenspace of  $X \otimes X$  orthogonal to  $\Pi_\pm \otimes P_\mp$  (also of dimension  $d_A^2/4$ ), and arbitrary rank- $k_1$  and  $r - k_1$  projectors  $\Pi_{++}$  and  $\Pi_{+-}$ . This is possible as long as  $k_2 \leq d_A^2/4$ ,  $d - r - k_2 \leq d_A^2/4$ ,  $k_1 \leq d_A^2/4$  and  $r - k_1 \leq d_A^2/4$ . Substituting our choice of  $r$ ,  $k_1$  and  $k_2$ , we find that it suffices to take  $\lceil d/4 \rceil \leq d_A^2/4$ . ■

Corollary 12 shows that, for suitable dimensions, a subspace chosen at random subject to the symmetry and orthogonality constraints of Eqs. (19c), (19d) and (19g) from Theorem 3 will, with probability 1, satisfy the strong unextendibility conditions of Eqs. (19a) and (19b). But Theorem 14 shows that

there is a non-zero probability that such a random subspace will satisfy the positivity conditions of Eqs. (19e) and (19f). Therefore, for suitable dimensions, there must exist at least one subspace  $S$  satisfying all the conditions of Theorem 3. Hence, by that theorem, there exists a pair of channels  $\mathcal{E}_{1,2}$  with  $C_0(\mathcal{E}_{1,2}) = 0$  but  $Q_0(\mathcal{E}_1 \otimes \mathcal{E}_2) \geq 1$ .

Satisfying all the dimension requirements of Corollary 12 and Theorem 14 imposes constraints on the channel input and output dimensions  $d_A$  and  $d_B$ , and number of Kraus operators  $d_E$  (which corresponds to the subspace dimension  $d$ ). Together, these constraints impose  $d_A \geq 48$  and  $d_E \geq 12(2d_A - 1)$ , giving our main result:

**Theorem 15** *Let  $d_A = 48$ ,  $d_E = 12(2d_A - 1) = 1140$  and  $d_B = d_A d_E = 54720$ . Then there exist channels  $\mathcal{E}_1, \mathcal{E}_2$  such that:*

- Each channel  $\mathcal{E}_{1,2}$  maps  $\mathbb{C}^{d_A}$  to  $\mathbb{C}^{d_B}$  and has  $d_E$  Kraus operators.
- Each channel  $\mathcal{E}_{1,2}$  has no classical zero-error capacity (hence no quantum zero-error capacity either).
- The joint channel  $\mathcal{E}_1 \otimes \mathcal{E}_2$  has positive quantum zero-error capacity (hence all other capacities are non-zero).

This trivially implies that there exist channels with similar properties in all dimensions larger than these, too.

## V. CONCLUSIONS

There has been a recent a surge of progress in the theory of quantum channels, especially their capacities. We now know that two uses of a quantum channel can sometimes, by using entangled signal states, transmit more than twice as much classical information as a single use [13]. This makes it likely that any expression for the classical capacity will require regularisation, implying that it cannot be computed in general. We have known for some time that this is also the case for the quantum capacity [14], but we now also know that the quantum capacity itself is non-additive. Indeed, it exhibits the particularly extreme form of non-additivity known as superactivation [4]. This implies that the amount of quantum information that can be sent through a channel depends on what other channels are also available. Understanding these additivity violations is now a key goal of quantum information theory.

Both manifestations of non-additivity—regularisation and non-additive capacity—are already displayed by the *zero-error* capacity of *classical* channels [1], [15], [16], though superactivation remains impossible even in the zero-error setting. Zero-error capacities have been the subject of intense study in the classical information theory literature for over half a century. They are therefore an interesting area in which to probe *quantum* channel capacities, and attempt to understand non-additivity phenomena. Non-additivity in the purely classical setting obviously has nothing to do with entanglement. But quantum channels display even stronger non-additivity than their classical counterparts. In the quantum world, the presence of entanglement *does* lead to superactivation of the classical zero-error capacity of quantum channels [5].

The usual classical and quantum capacities are not at all closely related. There is no reason to expect that channels



displaying additivity violations for the quantum capacity will possess any interesting additivity properties for the classical capacities, or vice versa. As a consequence, the recent non-additivity results for the usual capacities [4], [13] required very different mathematical techniques for the two cases.

However, in the zero-error setting, this work shows a striking non-additivity phenomenon that connects the classical and quantum capacities. We have proven the existence of pairs of channels that, individually, can not communicate *any* information with zero error, even classical information. But, when used together, even a single use of the joint channel suffices to communicate *all* forms of information, quantum and classical. These channels therefore exhibit the most extreme possible form of additivity violation; their zero-error capacities *simultaneously* violate additivity for both classical and quantum information, and in the most extreme way (superactivation) to boot. This extreme form of superactivation is trivially impossible for classical channels, or for the usual capacities of quantum channels. Zero-error communication therefore provides a compelling setting in which to explore non-additivity phenomena in quantum information theory.

#### REFERENCES

- [1] C. E. Shannon, "The zero-error capacity of a noisy channel," *IRE Trans. Inform. Theory*, vol. IT-2, p. 8, 1956.
- [2] J. Körner and A. Orlitsky, "Zero-error information theory," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, p. 2207, 1998.
- [3] R. A. C. Medeiros and F. M. de Assis, "Quantum zero-error capacity," *Int. J. Quant. Inf.*, vol. 3, p. 135, 2005.
- [4] G. Smith and J. Yard, "Quantum communication with zero-capacity channels," *Science*, vol. 321, p. 1812, 2008, (arXiv:0807.4935 [quant-ph]).
- [5] T. S. Cubitt, J. Chen, and A. W. Harrow, "Superactivation of the asymptotic zero-error classical capacity of a quantum channel," arXiv:0906.2547.
- [6] R. Duan, "Superactivation of zero-error capacity of noisy quantum channels," arXiv:0906.2527 [quant-ph], 2009.
- [7] J. Harris, *Algebraic Geometry*. Springer, 1992.
- [8] H. Barnum and E. Knill, "Reversing quantum dynamics with near-optimal quantum and classical fidelity," *J. Math. Phys.*, vol. 43, no. 5, pp. 2097–2106, 2002.
- [9] C. King and M. Ruskai, "Minimal entropy of states emerging from noisy quantum channels," *IEEE Trans. Info. Theory*, vol. 47, pp. 192 – 209, 2001.
- [10] T. Cubitt, A. W. Harrow, D. Leung, A. Montanaro, and A. Winter, "Counterexamples to additivity of minimum output p-rényi entropy for p close to 0," *CMP*, vol. 284, p. 281, 2008, (arXiv:0712.3628 [quant-ph]).
- [11] R. Duan, J. Chen, and Y. Xin, "Unambiguous and zero-error classical capacity of noisy quantum channels," (Manuscript in preparation).
- [12] R. Bhat, "A completely entangled subspace of maximal dimension," *Int. J. Quant. Inf.*, vol. 4, no. 2, p. 325, 2006.
- [13] M. B. Hastings, "A counterexample to additivity of minimum output entropy," *Nature Physics*, vol. 5, 2009, (arXiv:0809.3972 [quant-ph]).
- [14] D. P. DiVincenzo, P. W. Shor, and J. A. Smolin, "Quantum channel capacity of very noisy channels," *PRA*, vol. 57, p. 830, 1998, (arXiv:quant-ph/9706061).
- [15] W. Haemers, "On some problems of lovasz concerning the shannon capacity of a graph," *IEEE Trans. Inform. Theory*, vol. 25, pp. 231–232, 1979.
- [16] N. Alon, "The shannon capacity of a union," *Combinatorica*, vol. 18, no. 3, pp. 301–310, 1998.